



Available Online at ESci Journals International Journal of Phytopathology

ISSN: 2305-106X (Online), 2306-1650 (Print)
<http://www.escijournals.net/phytopathology>



MOLECULAR CHARACTERIZATION OF THE MOVEMENT AND COAT PROTEINS OF A NEW ELM MOTTLE VIRUS ISOLATE INFECTING EUROPEAN WHITE ELM (*ULMUS LAEVIS* PALL.)

Anne-Mareen Eisold

Department of Biology, Institute of Life Sciences, Humboldt-Universität zu Berlin, D - 10115 Berlin, Germany.

ABSTRACT

European white elms (*Ulmus laevis* Pall.) growing in a park in Caputh near Berlin (Germany) were regularly monitored over a period of 18 years showing virus infection-like symptoms such as chloroses, chlorotic ringspots, mottling and dieback. To obtain the evidence for viral infection, RNA-seq using an Illumina Hi Seq2500 was conducted and three contigs were obtained. They match with the three EMoV genomic RNAs and cover the open reading frames for the viral replicase, the polymerase and the movement and coat proteins (MP, CP). The contigs show identities of 95.3–96.4%, 91.9–93.3% and 89.0–92.5% at the nucleotide level with RNA 1, RNA 2 and RNA 3 of reference sequences, respectively. The analyses of the MP and CP showed significant differences in amino acid sequence compositions compared to those of reference EMoV sequences. These results demonstrate the presence of a so far unknown isolate of EMoV. This is the first report of sequence data of EMoV infecting *U. laevis*.

Keywords: Illumina high throughput sequencing, EMoV, RNA 3 variability, sequence analysis

INTRODUCTION

A population of European white elms (*Ulmus laevis* Pall.) in the park in Caputh (federal state Brandenburg, Germany) suffer from chlorotic ringspots, mottling and necrosis since decades (Figure 1). Assuming a virus to be the causal agent, long-term investigations were conducted and revealed an increase of the extent of symptoms on the affected branches. At the same time, distinct dieback and a decline of the annual growth was observed. Investigations in identification and characterization of the causal pathogen have been conducted for several years. In previous investigations viruses known to infect elm trees such as *Elm mottle virus* (EMoV) were excluded as the causal agent applying biological and serological methods (Bandte *et al.*, 2004). Nonetheless, recent findings reveal the presence of EMoV in German elms with described symptoms (Büttner *et al.*, 2015). Although, the EMoV belonging to subgroup 2 of the genus *Ilarvirus* (*Bromoviridae*) was recently described as common in elm species (EPPO, 2017), the number of recorded detections is rare (Table 1).

EMoV has a tripartite positive single-stranded RNA ((+) ssRNA) genome with cap structures at the 5'-termini (Figure 2). The three RNA molecules lack polyadenylations at the highly conserved 3'-termini but form strong non-aminoacylated secondary structures. The EMoV genomic RNAs code for four open reading frames (ORFs) and are separately embedded forming quasi-isometric and bacilliform virions, respectively (King *et al.*, 2012). RNA 1 (3,431 bp) codes for the replicase (ORF 1) comprising the conserved domains (CDs) for the viral methyltransferase (Vmethyltransf, pfam1660) and a helicase motif (Viral_helicase1, pfam1443). This and the RNA-dependent RNA polymerase (RdRP_2, pfam00978) coding within ORF 2 of the RNA 2 (2,874 bp) are regarded to be subunits of the viral replicase complex with strictly linked and coordinated RNA 1 and RNA 2 replication strategies (Bol, 2005). The RNA 2 of EMoV includes a second ORF (ORF 2b) that codes for a 2b protein that has been proven to be involved in gene silencing (Shimura *et al.*, 2013). RNA 3 (2,325 bp) encodes the MP (proximal ORF 3a) belonging to the "30K" Bromo_MP superfamily (Melcher, 2000; pfam01573) and the CP (distal ORF 3b), which comprises the Ilar_coat superfamily CD (pfam01787).

* Corresponding Author:

Email: anne-mareen.eisold@hu-berlin.de

© 2019 ESci Journals Publishing. All rights reserved.

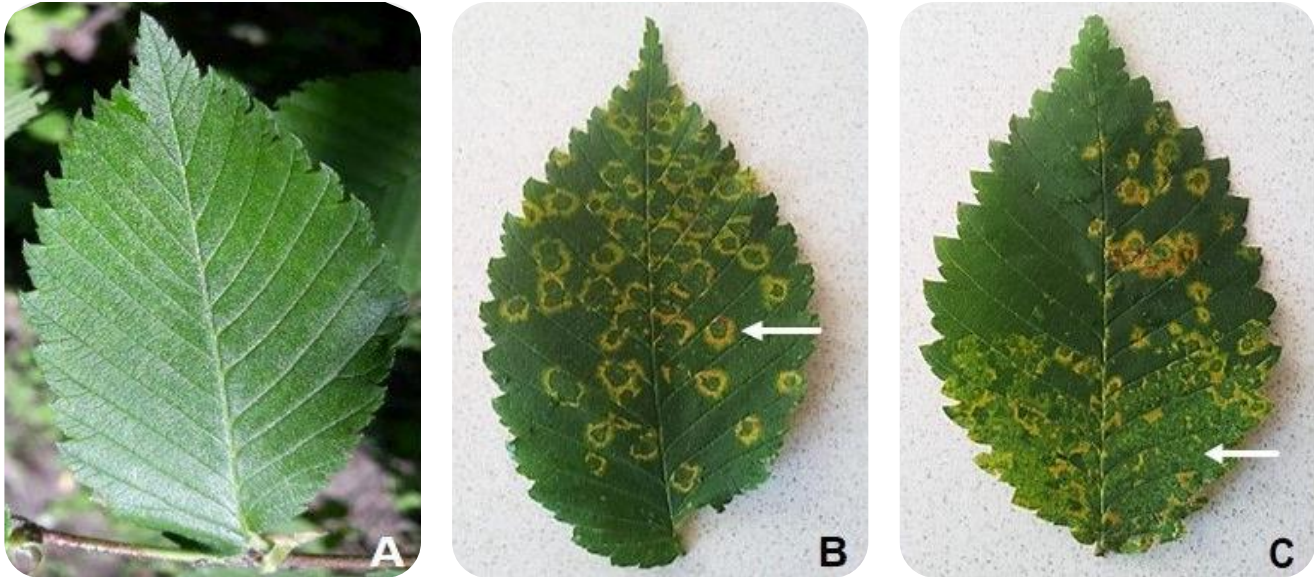


Figure 1. Symptoms on leaves of *U. laevis*; (A) leaf without symptoms, (B) leaf showing ringspots and (C) mottling.

Table 1. Natural host range and sequence accessions of EMoV isolates available in NCBI database. Asterisks indicate references providing sequence accessions.

| Natural host | Geographical origin | References | # Accession |
|------------------------------|--------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Hydrangea macrophylla</i> | Great Britain | (Thomas <i>et al.</i> , 1983; Scott <i>et al.</i> , 2003*) | RNA 1: AH009667.2, AH009668.2 RNA 2: AH009666.2 RNA 3: AF172965.1 |
| <i>Philadelphus</i> sp. | Germany | (Schmelzer, 1974) | |
| <i>Syringa vulgaris</i> | Germany | (Schmelzer <i>et al.</i> , 1966) | |
| <i>Ulmus glabra</i> | Germany Great Britain | (Schmelzer <i>et al.</i> , 1966; Schmelzer, 1969; Jones and Mayo, 1973; Ge <i>et al.</i> , 1997*; Scott <i>et al.</i> , 2003*; Untiveros <i>et al.</i> , 2010*) | RNA 1: U57047.1, U85402.1, U85401.1, NC_003569.1 RNA 2: U34050.1, U85400.1, NC_003568.1, GQ865661.1 RNA 3: U57048.1, U85399.1, NC_003570.1 |
| <i>Ulmus minor</i> | Croatia Germany | (Schmelzer <i>et al.</i> , 1966; Schmelzer, 1969; Pleše and Juretić, 1999) | |
| <i>Ulmus laevis</i> | Germany | (Büttner <i>et al.</i> , 2015) this work* | RNA 1: LT898351.1 RNA 2: LS422792.1 RNA 3: LT898352.1 |
| <i>Ulmus</i> sp. | Bulgaria Czech Republik Russia Slovakia | (Schmelzer <i>et al.</i> , 1966) | |

ORF 3b is expressed by a subgenomic RNA 4 (sgRNA) (King *et al.*, 2012) using an internal promotor with specific secondary structures, which is presumed to be a prerequisite for sgRNA synthesis (Jaspars, 1998; Aparicio and Pallás, 2002).

This work shows sequence analysis of EMoV genomic RNAs. It highlights the analysis of the EMoV RNA 3, coding for the viral MP and CP and provides new insights into the variability of the RNA 3 on nucleotide and amino acid levels. Data show an as yet unreported variability in

sequence compositions and point to the presence of a new EMoV strain infecting *U. laevis* in Germany.

MATERIALS AND METHODS

In 2014, leaves of a diseased tree were sampled. Complementary to the EMoV-specific RT-PCR described by Büttner *et al.* (2015) a next generation sequencing method was employed. Areas with ringspots were cut from symptomatic leaves and 70 mg fresh material was used for RNA extraction with Invitrap Spin Plant RNA Mini Kit (STRATEC Molecular). The NucleoSpin® RNA Kit (Macherey-Nagel) containing rDNase was used for removing residual DNA, followed by cleaning the sample with NucleoSpin® RNA clean-up Kit (Macherey-Nagel). Efficient depletion of plant large ribosomal RNA from total RNA was accomplished using a RiboMinus Plant Kit for RNA-Sequ (Invitrogen) and 10 µg of high integrity total RNA. Double-stranded full-length cDNA was synthesized with 1-2 µg RiboMinus RNA using the Maxima H Minus Double-Stranded cDNA Synthesis Kit (Thermo Scientific) primed with random hexamers. All kits were executed according to manufacturer's instructions. Approximately 1-2 µg of double-stranded cDNA were sent to BaseClear (Netherlands) for RNA sequence analysis. Paired-end 100 bp sequence reads (≈ 50 Mb) were generated conducting Illumina Hi Seq2500 system. The reads were mapped, and virus sequences *de novo* assembled on Biolinux and CLC Genomics Workbench, respectively. Out of a dataset of 1,011,396 paired-end reads, 908 contigs were constructed and used to identify viral sequences, which were analyzed with Clustal W (Larkin *et al.*, 2007) and Geneious version 9.1.3 (Kearse *et al.*, 2012). Genomic EMoV RNA was identified with BLASTX 2.2.25 (Altschul *et al.*, 1990). The resulting EMoV RNA sequences were aligned with reference EMoV isolates in order to determine variability in sequence composition. ORF 3a and 3b were verified conducting a PCR followed by Sanger sequencing (data not shown). Partial sequences of RNA 1, RNA 2 and RNA 3 of the new EMoV isolate tentatively named 'Berlin' were deposited at European Nucleotide Archive (ENA) under accession numbers LT898351, LS422792 and LT898352. Search for protein domains was conducted with Conserved Domain Database at NCBI (Marchler-Bauer *et al.*, 2008), protein sequences were computed using Geneious version 9.1.3.

RESULTS AND DISCUSSION

RNA deep sequencing resulted in three contigs that are 3,210 nt, 2,542 nt and 2,175 nt in length and match with EMoV reference sequences available in the NCBI

database. The contigs were confirmed to be partial sequences of EMoV RNA 1, RNA 2 and RNA 3 comprising ORFs coding for the replicase (ORF 1), polymerase (ORF 2a), MP (ORF 3a) and CP (ORF 3b) (Figure 2).

RNA 1 shows identities between 95.3% and 96.4% and within RNA 2 between 91.9% and 93.3% of the isolate 'Berlin' and reference sequences obtained from *U. glabra* and *H. macrophylla* (Table 1). These coincide in 99.3 – 100% within RNA 1 and in 98.2 – 100% within RNA 2 to each other. Thereby, sequences of RNA 1 and RNA 2 of the German EMoV isolate differ distinctly from reference sequences. The partial sequence of RNA 3 of the EMoV isolate 'Berlin' covers the 5'-non translated region (NTR), ORF 3a encompassing the interval 364..1,215, the non-coding intergenic region, ORF 3b encompassing the interval 1,344..1,997 and the partial 3'-NTR. Reference nucleotide sequences of EMoV RNA 3 (U57048.1, U85399.1, NC_003570.1 and AF172965.1) that have been isolated from *U. glabra* and *H. macrophylla* plants in Great Britain (Thomas *et al.*, 1983; Ge *et al.*, 1997; Scott *et al.*, 2003), show 89.0 – 92.5% identity with the German EMoV isolate from *U. laevis* (Table 2). This value is remarkably lower than the identities between the British isolates that show between 95.0% and 100% identity.

ORF 3a is 852 nt long with a computed weight of 31.3 kDa. The sequence identity between the British isolates ranges between 95% and 99.7% whereas the identity to the German isolate is 93.7% and 94.1% respectively. Within ORF 3a, 49 base substitutions were found. On the amino acid level six substitutions occur by replacing F₆₇ with L₆₇, I₇₅ with L₇₅, R₁₅₇ with K₁₅₇, P₂₅₈ with S₂₅₈, R₂₆₀ with G₂₆₀ and H₂₈₀ with R₂₈₀. Therefore, the protein sequence of the German isolate share 97.5% identity with the reference sequences, while they show identities between 99.3% and 100% to each other (Table 2). The MP of the German isolate shows six amino acid substitutions from which residues 67 and 75 are located within one of the two conserved RNA-binding domains (RBD). Hybridization studies revealed that some basic amino acids in RBD are essential for the viral capability for cell-to-cell movement as well as for RNA-binding (Herranz *et al.*, 2005). In the case of EMoV isolate 'Berlin', the residues R₆₂, R₇₃ and K₇₄ are basic and residues K₇₁ and H₇₉ are highly conserved (Pallás *et al.*, 2012). By substituting F₆₇ and I₇₅ with L that all are non-polar a change within the RBD motif polarity does not occur. Accordingly, the substitutions within the isolate 'Berlin' are not assumed neither to hinder RNA affinity nor cell-to-cell movement.

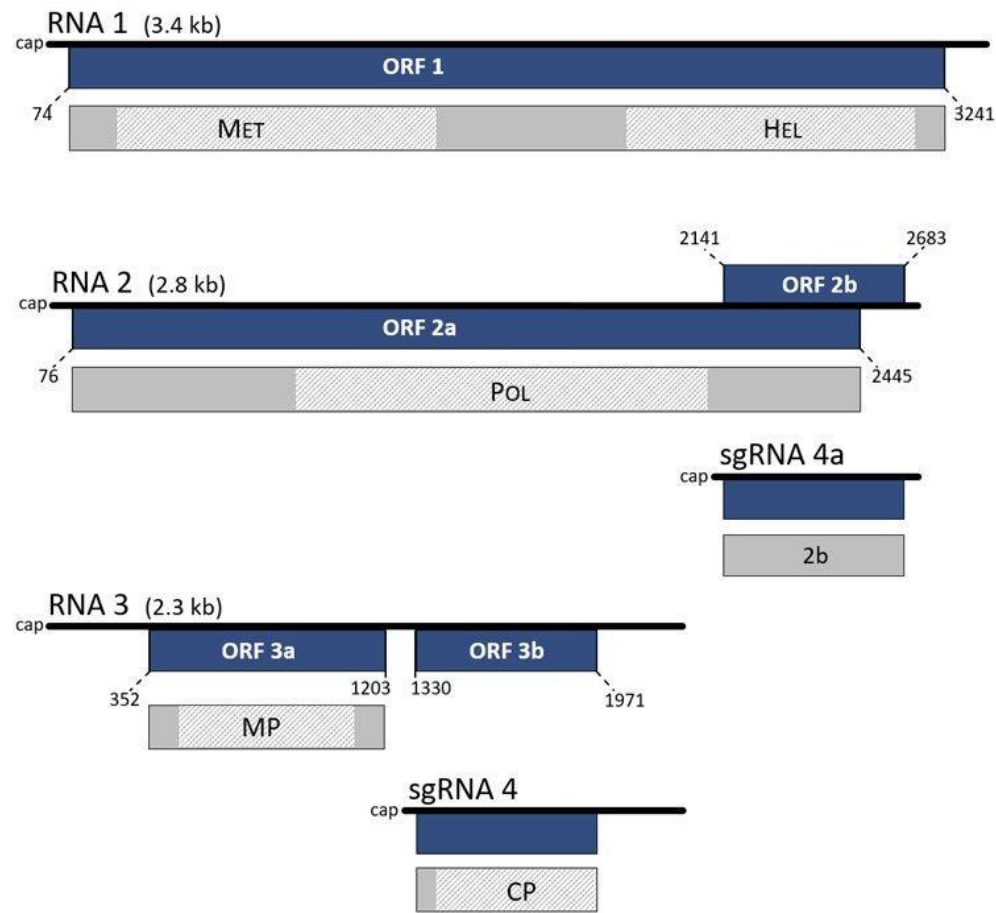


Figure 2. Genome organization of *Elm mottle virus*; genomic RNAs 1, 2 and 3 as well as subgenomic RNAs (sgRNA) 4 and 4a are visualized as blue boxes for open reading frames (ORF); viral proteins are represented by grey boxes and characteristic domains methyltransferase (Met), helicase (Hel), RNA-dependent RNA polymerase (Pol), movement protein (MP) and coat protein (CP) are indexed in light grey; cap structures at 5'-terminus are indicated; locations of ORF intervals are shown relative to EMoV sequences U57047.1 (RNA 1), U34050.1 (RNA 2), U57048.1 (RNA 3).

Table 2. Comparison of RNA 3 of EMoV isolate from *U. laevis* (LT898352.1) with reference isolates obtained from *U. glabra* and *H. macrophylla* respectively. Identity matrix shows percentages of identities on nucleotide (nt) level of total RNA 3 and on amino acid level of movement protein (MP) and coat protein (CP).

| Accession/ Host | U57048.1 | | U85399.1 | | NC_003570.1 | | AF172965.1 | | LT898352.1 | |
|-------------------------------------|----------|---------|----------|------------|-------------|------------|------------|----------|------------|-------------|
| | nt | MP / CP | nt | MP / CP | nt | MP / CP | nt | MP / CP | nt | MP / CP |
| U57048.1 <i>U. glabra</i> | | | | 99.3/ 98.6 | | 99.3/ 98.6 | | 99.3 / - | | 97.5 / 89.4 |
| U85399.1 <i>U. glabra</i> | 95.0 | | | | | 100 / 100 | | 99.3 / - | | 97.5 / 88.9 |
| NC_003570.1 <i>U. glabra</i> | 95.1 | | 100 | | | | | 99.3 / - | | 97.5 / 88.9 |
| AF172965.1 <i>H. macrophylla</i> | 95.2 | | 99.7 | | 99.7 | | | | | 97.5 / - |
| LT898352.1 <i>U. laevis</i> | 89.0 | | 92.5 | | 92.5 | | 92.5 | | | |

The nucleotide sequence of ORF 3b show 57 base substitutions. Additionally, the sequence obtained from *U. laevis* possess three prominent nucleotide insertions in ORF 3b. At positions 1,469 and 1,499, two nucleotide triplets were found comprising a GGG and a GCA motif. Additionally, an insertion of a CACAAA motif augment the open reading frame at position 1,526 (Figure 3A). These insertions do not shift the open reading frame, but they are translated as additional amino acids (Figure 3B). Consequently, the computed 24.0 kDa CP of the EMoV isolate 'Berlin' contains 217 instead of 213 amino acids. The additional amino acids glycine, glutamine, asparagine and proline characterize the CP structure of this EMoV isolate and distinguish it significantly from those of isolates obtained from *U. glabra* and *H. macrophylla*. Within the first 25 amino acids of the N-terminus of the CP that are proven to be mandatory for binding to viral RNA (Bol, 2005), the German isolate differs at three positions from reference isolates by replacing A₁₆ with T₁₆ and substituting G₂₃ with R₂₃ and S₂₄ with G₂₄. Overall, 17 out of 217 amino acids are uniquely substituted within the CP of the EMoV isolate 'Berlin'. To initiate infection of ilarviruses the binding of CP to the 3'-NTR of the genomic RNAs is required (phenomenon of genome

activation). Thus, for the initiation as well as for the propagation of the viral infection, the CP acts as structural key component (Sánchez-Navarro *et al.*, 1997; Pallás *et al.*, 1999; Bol, 2005; MacFarlane and McGavin, 2009). Within the N-terminus of the EMoV CP, an arginine-rich basic motif that is considered to bind to the 3'-NTR of its RNA 3 is determined between amino acid residues 22 and 44 (Pallás *et al.*, 1999; Aparicio *et al.*, 2003; Pallás *et al.*, 2013). The R₃₄ is proposed to be the central essential residue for specific binding of the CP to the 3'-NTR (Ansel-McKinney *et al.*, 1996).

Within this motif, the EMoV isolate 'Berlin' has three substitutions and one insertion (Figure 3A) augmenting the number of arginine residues. This leads to the assumption, that R₂₅, R₂₈, R₂₉, R₃₄ and R₃₆ are conserved within the EMoV CP, whether R₂₃ appears only in the German isolate. With regard to the functional role of the basic motif for the infection process, the content of six instead of five arginine residues is supposed to support genome activation process. The scientific proof of the question, if the higher content of arginine within this region has an enhancing impact to the infectivity of this isolate, remains for further investigations, as well as a putative host specificity.

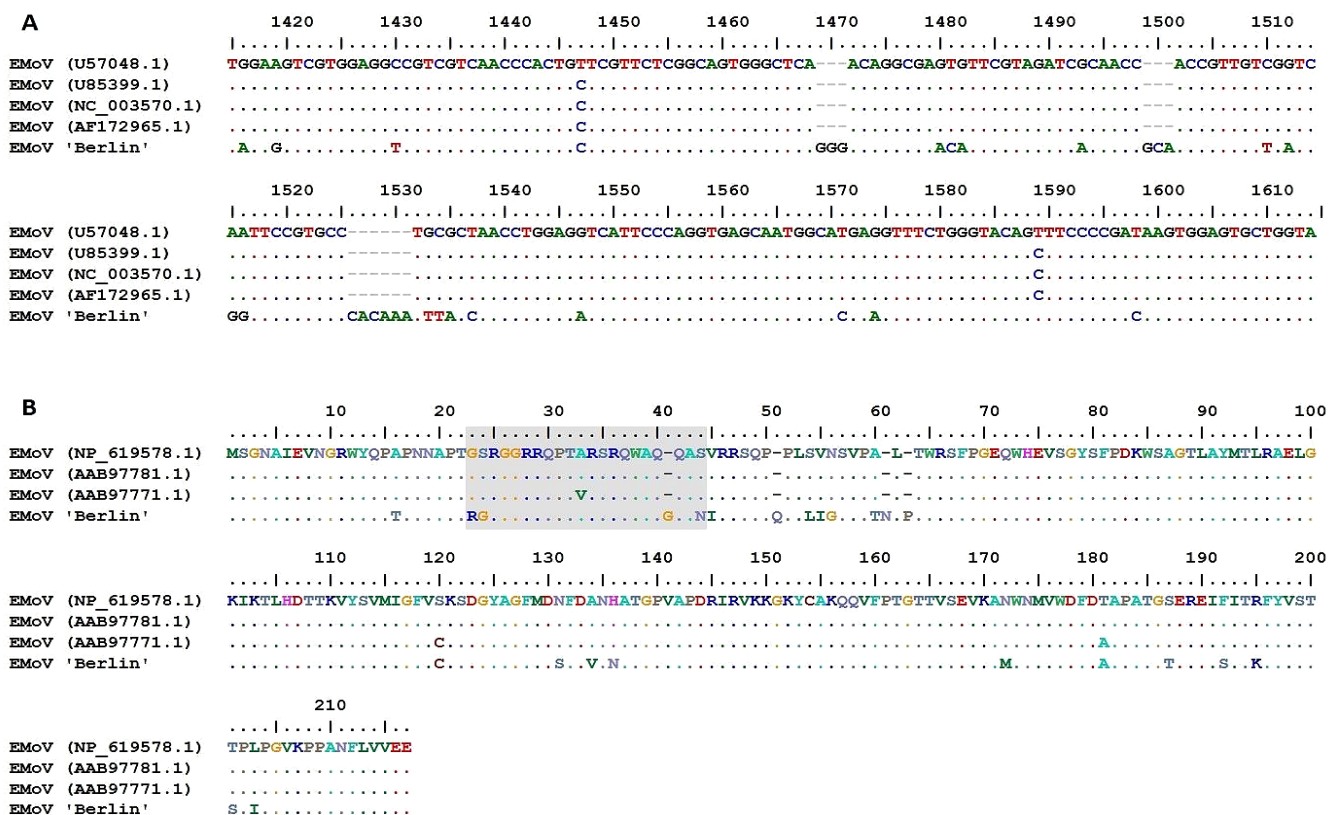


Figure 3. Alignment of EMoV coat protein nucleotide (A, partially shown) and amino acid (B) sequences; the arginine rich motif is labelled in gray. Distinct letters mark sequence variations, identities are labelled as dots, gaps are labelled as bars.

The EMoV isolate investigated here features distinct differences at nucleotide as well as at amino acid sequence that distinguish it remarkably from other EMoV isolates. This study contributes in data collection and genome analysis of the EMoV augmenting the knowledge about viruses infecting elms. It provides for the first time sequence data of a so far unknown EMoV isolate affecting *U. laevis*.

ACKNOWLEDGEMENT

The author gratefully thanks the Division Phytomedicine at the Humboldt-Universität zu Berlin, in particular Martina Bandte for contributing to the sample collection, Markus Rott for supporting the preparation steps, Artemis Rumbou for establishing the NGS project and Carmen Büttner for facilitating the experiments by providing the laboratory equipment and funding of the Illumina high throughput sequencing technique.

COMPLIANCE WITH ETHICAL STANDARDS

The author declares that ethical standards have been followed and that no human participants or animals were involved in this research.

COMPETING INTERESTS

The author declares that there are no competing interests.

REFERENCES

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215: 403-10.
- Ansel-McKinney, P., S. W. Scott, M. Swanson, X. Ge and L. Gehrke. 1996. A plant viral coat protein RNA binding consensus sequence contains a crucial arginine. *The EMBO Journal*, 15: 5077-84.
- Aparicio, F. and V. Pallás. 2002. The molecular variability analysis of the RNA 3 of fifteen isolates of prunus necrotic ringspot virus sheds light on the minimal requirements for the synthesis of its subgenomic RNA. *Virus Genes*, 25: 75-84.
- Aparicio, F., M. Vilar, E. Perez-Payá and V. Pallás. 2003. The coat protein of prunus necrotic ringspot virus specifically binds to and regulates the conformation of its genomic RNA. *Virology*, 313: 213-23.
- Bandte, M., M. Essing, C. Obermeier and C. Büttner. 2004. Virus-diseased *Ulmus laevis* in eastern Germany. *Forest Systems*, 13: 65-69.
- Bol, J. F. 2005. Replication of alfamo and ilar viruses: Role of the coat protein. *Annual Review of Phytopathology*, 43: 39-62.
- Büttner, C., S. von Bargaen, A.-M. Eisold, M. Bandte and M. Rott. 2015. Eine Fallstudie zum Elm mottle virus (EMoV) an Ulme (*Ulmus* sp.). In: D. Dujesiefken (ed.), *Jahrbuch der Baumpflege* Braunschweig: Haymarket Media.
- EPPO. 2017. PQR database European and Mediterranean Plant Protection Organization. Paris, France.
- Ge, X., S. W. Scott and M. T. Zimmerman. 1997. The complete sequence of the genomic RNAs of spinach latent virus. *Archives of Virology*, 142: 1213-26.
- Herranz, M. C., J. A. Sánchez-Navarro, A. Saurí, I. Mingarro and V. Pallás. 2005. Mutational analysis of the RNA-binding domain of the Prunus necrotic ringspot virus (PNRSV) movement protein reveals its requirement for cell-to-cell movement. *Virology*, 339: 31-41.
- Jaspars, E. M. J. 1998. A core promoter hairpin is essential for subgenomic RNA synthesis in alfalfa mosaic alfamovirus and is conserved in other Bromoviridae. *Virus Genes*, 17: 233-42.
- Jones, A. T. and M. A. Mayo. 1973. Purification and properties of elm mottle virus. *Annals of Applied Biology*, 75: 347-57.
- Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes and A. Drummond. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28: 1647-49.
- King, A. M. Q., M. J. Adams, E. B. Carstens and E. J. Lefkowitz. 2012. *Virus taxonomy: Classification and nomenclature of Viruses Ninth report of the International Committee on Taxonomy of Viruses*. Academic Press Elsevier. San Diego, USA. pp. 486-87.
- Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23: 2947-48.
- MacFarlane, S. A. and W. J. McGavin. 2009. Genome activation by raspberry bushy dwarf virus coat protein. *Journal of General Virology*, 90: 747-53.
- Marchler-Bauer, A., J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales and M. Gwadz. 2008. CDD: Specific functional annotation with the conserved

- domain database. Nucleic Acids Research, 37: 205-10.
- Pallás, V., F. Aparicio, M. C. Herranz, K. Amari, M. A. Sánchez-Pina, A. Myrta and J. A. Sánchez-Navarro. 2012. Iarviruses of *Prunus* spp.: A continued concern for fruit trees. Phytopathology, 102: 1108-20.
- Pallás, V., F. Aparicio, M. C. Herranz, J. A. Sánchez-Navarro and S. W. Scott. 2013. The Molecular Biology of Iarviruses Advances in Virus Research. Elsevier. pp. 139-81.
- Pallás, V., J. A. Sánchez-Navarro and J. Díez. 1999. In vitro evidence for RNA binding properties of the coat protein of prunus necrotic ringspot iarvirus and their comparison to related and unrelated viruses. Archives of Virology, 144: 797-803.
- Pleše, N. and N. Juretić. 1999. Virus disease of field elm (*Ulmus minor* Mill.) in Croatia. Šumarski List, 123: 95-100.
- Schmelzer, K. 1969. Das Ulmenscheckungs-Virus. Journal of Phytopathology, 64: 39-67.
- Schmelzer, K. 1974. Investigations on viruses of ornamental and wild shrubs. 8th communication: New findings on *Forsythia*, *Hydrangea* and *Philadelphus* and viruses on *Rhamnus*, *Centaurea*, *Galvezia*, *Cistus*, *Forestiera*, *Abeliophyllum*, *Celastrus*, *Staphylea* and *Crambe*. ZBL Bact II, 129: 139-68.
- Schmelzer, K., H. E. Schmidt and H. B. Schmidt. 1966. Viruskrankheiten und virusverdächtige Erscheinungen an Forstgehölzen. Archiv für Forstwesen, 15: 107-20.
- Scott, S. W., M. T. Zimmerman and X. Ge. 2003. Viruses in subgroup 2 of the genus Iarvirus share both serological relationships and characteristics at the molecular level. Archives of Virology, 148: 2063-75.
- Shimura, H., C. Masuta, N. Yoshida, K. Sueda and M. Suzuki. 2013. The 2b protein of Asparagus virus 2 functions as an RNA silencing suppressor against systemic silencing to prove functional synteny with related cucumoviruses. Virology, 442: 180-88.
- Sánchez-Navarro, J. A., J. F. Bol, C. B. Reusken and V. Pallás. 1997. Replication of alfalfa mosaic virus RNA 3 with movement and coat protein genes replaced by corresponding genes of Prunus necrotic ringspot iarvirus. Journal of General Virology, 78: 3171-76.
- Thomas, B. J., R. J. Barton and A. Tuszyński. 1983. Hydrangea mosaic virus, a new iarvirus from *Hydrangea macrophylla* (Saxifragaceae). Annals of Applied Biology, 103: 261-70.
- Untiveros, M., Z. Perez-Egusquiza and G. Clover. 2010. PCR assays for the detection of members of the genus Iarvirus and family Bromoviridae. Journal of Virological Methods, 165: 97-104.

